Serial Number: 09/864,156 Filing Date: May 25, 2001

Title: TEXT CLASSIFICATION SYSTEM

Page 2 Dkt: 1546.003US1

## IN THE CLAIMS

The claims stand as follows.

1. (ALLOWED) In a system comprising perspectives and categories, each perspective including at least one category representative of that perspective, a computerized method for classifying at least one item across multiple perspectives, said computerized method comprising:

associating category features with each category, wherein each of said category features represents one of a plurality of tokens;

producing a category vector for each category, wherein each category vector includes a weight corresponding to each category feature, said weight indicative of a degree of association between said category feature and said category;

associating item features with each item, wherein each of said item features represents one of a plurality of tokens found in said item;

producing a feature vector for each item, wherein each feature vector includes said item features with a count corresponding to each item feature, said count indicative of the number of times said item feature appears in said item;

multiplying said category vector by said item vector to produce a plurality of category scores for each item; and

for each perspective, across multiple perspectives, classifying an item into a category provided said category score exceeds a predetermined threshold.

2. (ALLOWED) The computerized method of claim 1, wherein

the count includes at least one subcount indicative of the number of times the item feature appears in a particular region of the item, wherein the particular region of the item is a subset of the item taken as a whole; and

for first and second categories, in classifying the item into the second category, substituting the at least one subcount for the count of each item feature depends on whether the item was classified into the first category.

Filing Date: May 25, 2001

e: TEXT CLASSIFICATION SYSTEM

Page 3 Dkt: 1546.003US1

3. (ALLOWED) The computerized method of claim 1, wherein

for first and second categories, classifying the item into the second category depends at least in part on a determination of whether the item was classified into the first category.

4. (ALLOWED) The computerized method of claim 1, wherein the classifying further comprises:

comparing a category score for a first item in a first perspective with a category score for said first item in a second perspective; and

modifying the category score in the first perspective in response to the category score in the second perspective.

- 5. (Cancelled)
- 6. (ALLOWED) The computerized method of claim 4, wherein said first perspective is an ancestor of said second perspective.
- 7. (Cancelled)
- 8. (ALLOWED) The computerized method of claim 4, wherein the comparing further comprises: creating an ordered list of category scores for all categories in all perspectives; identifying an item associated with a highest category score in a first perspective; identifying said item in a second perspective; and decreasing the category score of said item in said second perspective.
- 9-10 (Cancelled)
- 11. (ALLOWED) The computerized method of claim 8, further comprising repeating the identifying and decreasing for every perspective.
- 12. (ALLOWED) The computerized method of claim 1, wherein the weight corresponding to said concept node feature is between -1 and 1.

TEXT CLASSIFICATION SYSTEM

## 13-44. (CANCELLED)

- (ALLOWED) The computerized method of claim 1, in which the multiplying includes using 45. at least one attribute of at least one category vector in determining whether to include a document feature of the feature vector in the multiplying.
- (ALLOWED) The computerized method of claim 45, in which the using at least one 46. attribute includes using at least one of:
- a Stop attribute to indicate whether a feature must constitute something other than a stop word to be included in the multiplying;
- a Case attribute to indicate whether a feature must match a letter case specification to be included in the multiplying;
- a Stemming attribute to indicate whether a feature includes stemmed word forms to be included in the multiplying; and
- a Learned attribute to indicate how a human-specified feature is to be included in the multiplying.
- (ALLOWED) The computerized method of claim 45, in which the perspective to which the 47. category relates determines the value of the at least one attribute.

48-92. (CANCELLED)